

Automated geographic atrophy segmentation for SD-OCT images using region-based C-V model via local similarity factor

Sijie Niu,^{1,2,5} Luis de Sisternes,^{2,5} Qiang Chen,^{1,6} Theodore Leng,^{4,7}
and Daniel L. Rubin^{2,4,8}

¹*School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China*

²*Department of Radiology, Stanford University, Stanford, CA 94305, USA*

³*Byers Eye Institute at Stanford, Stanford University School of Medicine, Palo Alto, CA 9430, USA*

⁴*Department of Medicine (Biomedical Informatics Research), Stanford University School of Medicine Stanford, CA 94305, USA*

⁵*These authors contributed equally to this manuscript*

⁶*chen2qiang@njust.edu.cn*

⁷*tedleng@stanford.edu*

⁸*drrubin@stanford.edu*

Abstract: Age-related macular degeneration (AMD) is the leading cause of blindness among elderly individuals. Geographic atrophy (GA) is a phenotypic manifestation of the advanced stages of non-exudative AMD. Determination of GA extent in SD-OCT scans allows the quantification of GA-related features, such as radius or area, which could be of important value to monitor AMD progression and possibly identify regions of future GA involvement. The purpose of this work is to develop an automated algorithm to segment GA regions in SD-OCT images. An *en face* GA fundus image is generated by averaging the axial intensity within an automatically detected sub-volume of the three dimensional SD-OCT data, where an initial coarse GA region is estimated by an iterative threshold segmentation method and an intensity profile set, and subsequently refined by a region-based Chan-Vese model with a local similarity factor. Two image data sets, consisting on 55 SD-OCT scans from twelve eyes in eight patients with GA and 56 SD-OCT scans from 56 eyes in 56 patients with GA, respectively, were utilized to quantitatively evaluate the automated segmentation algorithm. We compared results obtained by the proposed algorithm, manual segmentation by graders, a previously proposed method, and experimental commercial software. When compared to a manually determined gold standard, our algorithm presented a mean overlap ratio (OR) of 81.86% and 70% for the first and second data sets, respectively, while the previously proposed method OR was 72.60% and 65.88% for the first and second data sets, respectively, and the experimental commercial software OR was 62.40% for the second data set.

©2016 Optical Society of America

OCIS codes: (100.0100) Image processing; (110.4500) Optical coherence tomography; (170.4470) Ophthalmology.

References and links

1. S. Resnikoff, D. Pascolini, D. Etya'ale, I. Kocur, R. Pararajasegaram, G. P. Pokharel, and S. P. Mariotti, "Global data on visual impairment in the year 2002," *Bull. World Health Organ.* **82**(11), 844–851 (2004).
2. R. G. Sayegh, C. Simader, U. Scheschy, A. Montuoro, C. Kiss, S. Sacu, D. P. Kreil, C. Prunte, and U. Schmidt-Erfurth, "A Systematic Comparison of Spectral-Domain Optical Coherence Tomography and Fundus Autofluorescence in Patients with Geographic Atrophy," *Ophthalmology* **118**(9), 1844–1851 (2011).
3. M. Fleckenstein, S. Schmitz-Valckenberg, C. Adrion, I. Krämer, N. Eter, H. M. Helb, C. K. Brinkmann, P. Charbel Issa, U. Mansmann, and F. G. Holz, "Tracking Progression with Spectral-Domain Optical Coherence

- Tomography in Geographic Atrophy Caused by Age-Related Macular Degeneration,” *Invest. Ophthalmol. Vis. Sci.* **51**(8), 3846–3852 (2010).
4. J. J. Wang, E. Rohtchina, A. J. Lee, E. M. Chia, W. Smith, R. G. Cumming, and P. Mitchell, “Ten-year incidence and progression of age-related maculopathy: the blue Mountains Eye Study,” *Ophthalmology* **114**(1), 92–98 (2007).
 5. R. Klein, B. E. Klein, M. D. Knudtson, S. M. Meuer, M. Swift, and R. E. Gangnon, “Fifteen-year cumulative incidence of age-related macular degeneration: the Beaver Dam Eye Study,” *Ophthalmology* **114**(2), 253–262 (2007).
 6. H. Buch, N. V. Nielsen, T. Vinding, G. B. Jensen, J. U. Prause, and M. la Cour, “14-year incidence, progression, and visual morbidity of age-related maculopathy: the Copenhagen City Eye Study,” *Ophthalmology* **112**(5), 787–798 (2005).
 7. I. Bhutto and G. Luttj, “Understanding age-related macular degeneration (AMD): relationships between the photoreceptor/retinal pigment epithelium/Bruch’s membrane/choriocapillaris complex,” *Mol. Aspects Med.* **33**(4), 295–317 (2012).
 8. R. P. Nunes, G. Gregori, Z. Yehoshua, P. F. Stetson, W. Feuer, A. A. Moshfeghi, and P. J. Rosenfeld, “Predicting the progression of geographic atrophy in age-related macular degeneration with SD-OCT en face imaging of the outer retina,” *Ophthalmic Surg. Lasers Imaging Retina* **44**(4), 344–359 (2013).
 9. J. S. Sunness, N. M. Bressler, Y. Tian, J. Alexander, and C. A. Applegate, “Measuring geographic atrophy in advanced age-related macular degeneration,” *Invest. Ophthalmol. Vis. Sci.* **40**(8), 1761–1769 (1999).
 10. J. S. Sunness, J. Gonzalez-Baron, C. A. Applegate, N. M. Bressler, Y. Tian, B. Hawkins, Y. Barron, and A. Bergman, “Enlargement of atrophy and visual acuity loss in the geographic atrophy form of age-related macular degeneration,” *Ophthalmology* **106**(9), 1768–1779 (1999).
 11. A. S. Lindblad, P. C. Lloyd, T. E. Clemons, G. R. Gensler, F. L. Ferris 3rd, M. L. Klein, and J. R. Armstrong; Age-Related Eye Disease Study Research Group, “Change in area of geographic atrophy in the Age-Related Eye Disease Study: AREDS report number 26,” *Arch. Ophthalmol.* **127**(9), 1168–1174 (2009).
 12. F. G. Holz, A. Bindewald-Wittich, M. Fleckenstein, J. Dreyhaupt, H. P. Scholl, and S. Schmitz-Valckenberg; FAM-Study Group, “Progression of geographic atrophy and impact of fundus autofluorescence patterns in age-related macular degeneration,” *Am. J. Ophthalmol.* **143**(3), 463–472 (2007).
 13. Z. Yehoshua, P. J. Rosenfeld, G. Gregori, W. J. Feuer, M. Falcão, B. J. Lujan, and C. Puliafito, “Progression of Geographic Atrophy in Age-Related Macular Degeneration Imaged with Spectral Domain Optical Coherence Tomography,” *Ophthalmology* **118**(4), 679–686 (2011).
 14. B. J. Lujan, P. J. Rosenfeld, G. Gregori, F. Wang, R. W. Knighton, W. J. Feuer, and C. A. Puliafito, “Spectral domain optical coherence tomographic imaging of geographic atrophy,” *Ophthalmic Surg. Lasers Imaging* **40**(2), 96–101 (2009).
 15. S. J. Chiu, J. A. Izatt, R. V. O’Connell, K. P. Winter, C. A. Toth, and S. Farsiu, “Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images,” *Invest. Ophthalmol. Vis. Sci.* **53**(1), 53–61 (2012).
 16. F. A. Folgar, E. L. Yuan, M. B. Sevilla, S. J. Chiu, S. Farsiu, E. Y. Chew, and C. A. Toth; Age Related Eye Disease Study 2 Ancillary Spectral-Domain Optical Coherence Tomography Study Group, “Drusen volume and retinal pigment epithelium abnormal thinning volume predict 2-year progression of age-related macular degeneration,” *Ophthalmology* **123**(1), 39–50 (2016).
 17. Q. Chen, L. de Sisternes, T. Leng, L. Zheng, L. Kutzscher, and D. L. Rubin, “Semi-automatic geographic atrophy segmentation for SD-OCT images,” *Biomed. Opt. Express* **4**(12), 2729–2750 (2013).
 18. Z. Hu, G. G. Medioni, M. Hernandez, A. Hariri, X. Wu, and S. R. Sadda, “Segmentation of the geographic atrophy in spectral-domain optical coherence tomography and fundus autofluorescence images,” *Invest. Ophthalmol. Vis. Sci.* **54**(13), 8375–8383 (2013).
 19. G. Tschepnakis, B. Lujan, O. Martinez, G. Gregori, and P. J. Rosenfeld, “Geometric deformable model driven by CoCRFs: application to optical coherence tomography,” *Med Image Comput Comput Assist Interv* **11**(Pt 1), 883–891 (2008).
 20. C. M. Li, C. Kao, J. Gore, and Z. Ding, “Implicit active contours driven by local binary fitting energy,” in: *IEEE Conference on Computer Vision and Pattern Recognition* **16**, 1–7 (2007).
 21. K. H. Zhang, H. H. Song, and L. Zhang, “Active contours driven by local image fitting energy,” *Pattern Recognit.* **43**(4), 1199–1206 (2010).
 22. S. Liu and Y. Peng, “A local region-based Chan-Vese model for image segmentation,” *Pattern Recognit.* **45**(7), 2769–2779 (2012).
 23. Z. Ji, Y. Xia, Q. Sun, G. Cao, and Q. Chen, “Active contours driven by local likelihood image fitting energy for image segmentation,” *Inf. Sci.* **301**, 285–304 (2015).
 24. T. F. Chan and L. A. Vese, “Active contours without edges,” *IEEE Trans. Image Process.* **10**(2), 266–277 (2001).
 25. S. Jiao, R. Knighton, X. Huang, G. Gregori, and C. Puliafito, “Simultaneous acquisition of sectional and fundus ophthalmic images with spectral-domain optical coherence tomography,” *Opt. Express* **13**(2), 444–452 (2005).
 26. Z. Yehoshua, C. A. A. Garcia Filho, F. M. Penha, G. Gregori, P. F. Stetson, W. J. Feuer, and P. J. Rosenfeld, “Comparison of geographic atrophy measurements from the OCT fundus image and the sub-RPE slab image,” *Ophthalmic Surg. Lasers Imaging Retina* **44**(2), 127–132 (2013).
 27. N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (1995).
 28. S. Osher and J. A. Sethian, “Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulation,” *J. Comput. Phys.* **79**(1), 12–49 (1988).

29. L. de Sisternes, J. Hu, D. L. Rubin, and M. F. Marmor, "Localization of damage in progressive hydroxychloroquine retinopathy on and off the drug: inner versus outer retina, parafovea versus peripheral fovea," *Invest. Ophthalmol. Vis. Sci.* **56**(5), 3415–3426 (2015).
30. L. de Sisternes, N. Simon, R. Tibshirani, T. Leng, and D. L. Rubin, "Quantitative SD-OCT imaging biomarkers as indicators of age-related macular degeneration progression," *Invest. Ophthalmol. Vis. Sci.* **55**(11), 7093–7103 (2014).
31. Q. Chen, S. Niu, H. Shen, T. Leng, L. de Sisternes, and D. L. Rubin, "Restricted summed-area projection for geographic atrophy visualization in SD-OCT images," *Transl. Vis. Sci. Technol.* **4**(5), 2 (2015).
32. R. Gonzales and R. Woods, *Digital Image Processing* (Prentice Hall, 1992), Chapter 6.
33. M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision* (CL Engineering, 1993), Chapter 6.

1. Introduction

Age-related macular degeneration (AMD) is the most common cause of legal blindness among elderly individuals in developed countries [1,2]. AMD is a chronic, progressive disease with various phenotypic manifestations, different disease stages, and variable rates of progression [3]. Choroidal neovascularization (CNV) and Geographic atrophy (GA) characterize late stages of the disease [4–6]. The latter occurs in 20% of patients with preexisting clinical hallmarks of this degenerative disease [2, 5]. The characteristic appearance of GA results from the loss of the photoreceptor layer, retinal pigment epithelium (RPE), and choriocapillaris [7,8]. In most cases, GA first appears in the parafoveal location and progresses around the fovea and then through the fovea with concomitant loss of central visual acuity [9,10]. Characterization of retinal regions affected by GA is fundamental in the diagnosis of advanced AMD because it could help clinicians to objectively monitor AMD progression, possibly identify regions of future GA involvement, or make a treatment decision. However, this characterization depends directly on the accurate segmentation and quantification of GA-affected regions and their properties, such as the extent, radius or area of abnormalities. Apart from being extremely challenging and time consuming, manual or semi-automatic GA segmentation is subject to user-variability, yielding potentially important differences. Thus, techniques to reliably and automatically outline and quantify GA regions are becoming important in the diagnosis of advanced AMD and for predicting future expansion of GA.

The appearance and progression of GA has been extensively studied using reflectance fundus imaging [11] and auto-fluorescence imaging [12] (2D topographic imaging techniques). However, these imaging techniques can only provide 2D topographic maps showing the atrophic area, and they are unable to resolve retina structure axially. The development of spectral-domain optical coherence tomography (SD-OCT) permitted the differentiation of retinal structures in the axial direction, generating 3D representations of retinal reflectivity and allowing an additional characterization of GA [13]. Macular SD-OCT images, normally covering a region of 6×6 mm of the macula region, are now becoming standard in an ophthalmic clinical visit and they allow the quantification of volumetric features that were previously inaccessible to fundus photography or FAF, for example, thickness and volume of the full retina or restricted to a particular collection of layers, or volume of drusen, cysts or defects in the ellipsoid zone. SD-OCT has also been used to characterize morphologic alterations that appear within the atrophic area, as well as within the surrounding retinal tissues. These quantifications have proven to be closely associated with visual acuity, enable assessment of multiple retinal diseases and conditions, and even predict disease progression. A necessary step to generate these quantifications in eyes affected by GA and to evaluate volumetric retina behavior in the presence of GA is obtaining an accurate outline of GA in SD-OCT images, therefore motivating this work. GA regions are normally visualized in SD-OCT by considering an axial projection of the 3-D data (*en face* OCT fundus image), which permits direct visualization of GA throughout the macula as a bright region within the image, owing to the increased penetration of light into the choroid where atrophy occurs. Previous studies [2, 14] have evaluated the utility of SD-OCT for grading GA as compared with FAF images and reported that SD-OCT projection images can be used to successfully identify and measure areas of GA. Therefore, SD-OCT seems to be an

appropriate imaging modality for automatically characterizing size, location and progression of GA lesions.

Although SD-OCT has proven successful in identifying and measuring GA regions, automatic segmentation of SD-OCT images containing GA is a difficult challenge due to low contrast and complicated retinal pathological structure. Previous work [15] has been focused on the estimation of the intra-retinal layers in eyes with GA using graph theory and dynamic programming, and the segmentation results were later used to measure the thickness of RPE, which can be used as a biomarker of GA appearance. However, this work did not produce outlines describing the GA region, as this was not the main focus of this work. The study of volume measurement of RPE as predictors of progression to advanced AMD has been reported by Folgar et al [16]. They discussed that the RPE-drusen complex abnormal thinning is believed to be an early precursor to the formation of GA lesions as well as a biomarker for the presence of any noncentral GA and the development of central GA over 2-year follow-up. Although the volume or thickness of RPE can be measured in SD-OCT images and considered as a biomarker of GA appearance, developing an automated segmentation algorithm solely based on the detection of missing regions of RPE is not straightforward. The relatively thin bright band representing the RPE in SD-OCT images, the increased signal underneath the regions of RPE loss and the characteristic noise in SD-OCT images makes the task of directly identifying GA regions by characterizing RPE loss in an automated fashion prone to error.

To our knowledge, very few methods [17–19] have been described for segmentation of GA in SD-OCT images. Those that have been reported provided good agreement between their methods and manually-defined GA regions, although still subject to several initialization limitations. These previous methods relied on the projection of the 3D OCT data in 2D *en face* images, which were later used to segment GA locations using level set functions. The limitations of these previous methods [17,18] is that an interactive contour selection around the GA region is needed to initialize a level set function, resulting in possible reader variability and complications when analyzing a large data set. A geometric deformable model driven by dynamically updated probability fields proposed by Tsechpenakis *et al.* [19] was employed to segment GA from *en face* SD-OCT images, and provided good segmentation performance. However, this model needed to use different markers for model initialization.

Segmentation models based on active contours and analysis of global or local image regions [20–24] have been widely studied and used due to their ability to adaptively handle the changes of topological structure, especially the Chan-Vese model (C-V) [24]. Although global region-based models [20, 24] have shown accurate image segmentation by using the statistical information inside and outside a contour to guide the curve evolution, their performance drops in images having intensity inhomogeneity. Later, local region-based models [21–23] were proposed to improve the segmentation performance on images with inhomogeneity. However, all of the region-based models need a markers selection for initialization. This fact and the presence of noise of different characteristics possibly influence their performance.

The purpose of this work is to develop a reliable, automatic, and effective approach for GA segmentation and quantification in SD-OCT images. Differently from the state-of-art local region-based models, the two intensity average functions in the improved C-V model are still global mean intensity rather than local mean intensity on the two sides of active contour, so the improved C-V model can be considered as a global region-based model. Moreover, the integration of the spatial distance within a local window and the difference between neighboring pixels and the global mean intensity is embedded in the C-V model to penalize the global data term so that the improved C-V model can reduce the influence of noisy pixels, which is called the local similarity factor. Two novelties of our method come from the construction of an automated initialization method and the development of an improved Chan-Vese model via local similarity factor (CVLSF) to generate the segmented outlines of GA-affected regions. The automatic initialization procedure includes an iterative threshold segmentation method and refinement of candidate regions using an intensity profile

set. The improvements made to the traditional Chan-Vese method [24] consider the introduction of a local similarity factor (LSF), which balances similarities observed by the spatial distance and gray level differences within a local window. The improved CVLSF model provides properties that allow the results to be less sensitive to noise. We conducted an extensive evaluation to demonstrate the ability of the proposed method segmenting GA in SD-OCT images acquired during clinical practice. We also evaluated our method's robustness to noise and compared it to that of a traditional Chan-Vese [24] approach in a synthetic image with different noise characterizations at several noise strength levels.

2. Methods

We have developed a fully automated pipeline for GA segmentation, as shown in Fig. 1. The data input comprises the series of SD-OCT scan data. The axial location of the layered structure in the SD-OCT scans is estimated using an intra-retinal segmentation algorithm [29], the results of which are used to generate topographic GA projection images [17]. We segment the coarse GA regions using an iterative segmentation method and then fill the missing regions with a set of GA candidate regions, extracted from an intensity profile set recorded at each horizontal location in each B-scan image. These results are then taken as the initialization for a modified region-based Chan-Vese (C-V) [24] method with local similarity factor (CVLSF), built to further identify and refine GA regions.

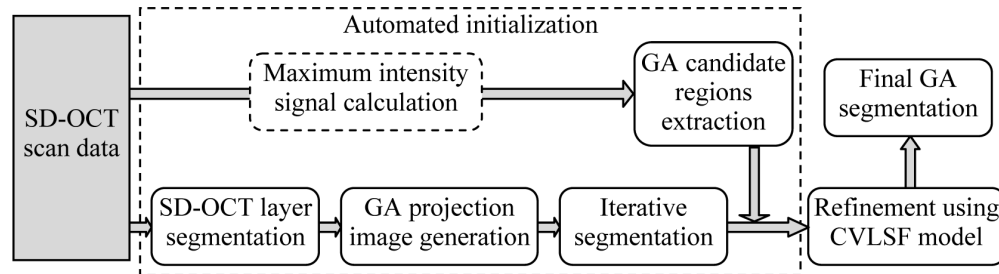


Fig. 1. The pipeline of the proposed automatic GA segmentation method.

2.1 Generation of GA projection image

There are three main methods for generating GA projection images from SD-OCT volumetric data sets: (1) The summed-voxel projection (SVP) [25], where the images are generated by adding all the voxel values in the 3D data along the axial A-scan lines; (2) the sub-RPE slab projection, available as the Cirrus Advanced RPE Analysis software in the Cirrus HD-OCT review software (version 6.0) [26] and the restricted summed-voxel projection method (RSVP) [17], both formed by projecting a region beneath the robust RPE fit in the axial direction; and (3) the restricted summed-area projection (RSAP) [31], which restricts the axial projection of SD-OCT images to the regions beneath the Bruch's membrane (BM) and also considers the choroidal vasculature's influence. The work presented here employs an approach similar to the sub-RPE slab method to generate a GA projection image, but restricting the averaged sub-volume to an upper limit indicated by a surface 39 microns (20 pixels in the images used in this work) below a fitted RPE outer boundary (estimated by automated segmentation [29]) and a lower limit h pixels below the upper limit surface. The parameter h indicates the maximum depth employed to generate the projection images and was set experimentally with the goal of maximizing GA visualization (see later experimental results and analysis section). Figure 2 shows an example comparing a traditional SVP projection image and the approach employed here in a patient with GA. We can observe how this restriction in the projected sub-volume increases the contrast observed between GA region and background. This restriction may also reduce the influence of other possible abnormalities like drusen or inner retina deformities in the projected image.

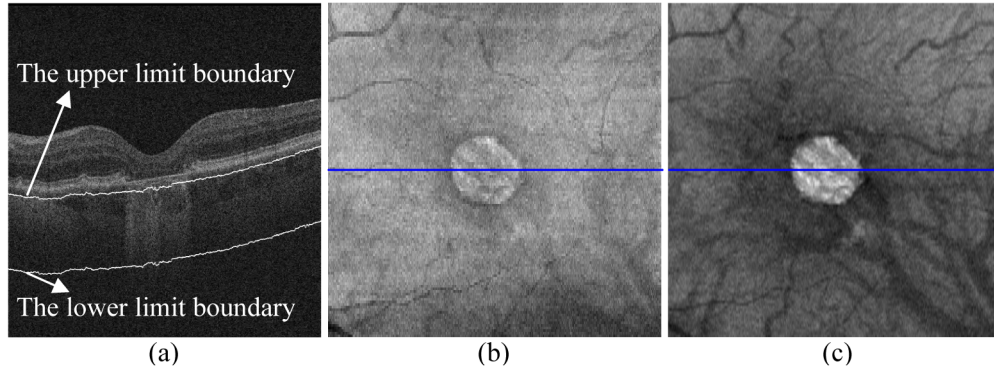


Fig. 2. An example of GA projection image generation. (a) B-scan across the center of the fovea, with the projection upper and lower limit boundaries used in our approach marked with white lines. (b) Traditional SVP GA projection image. (c) Resulting GA projection image used in our approach. The blue lines indicated in (b) and (c) correspond to the location of the B-scan across the center of the fovea shown in (a).

2.2 Iterative GA segmentation

Considering the intensity inhomogeneity and high noise level typically present in GA projection images (shown in Fig. 3(a) and 3(b)), conventional thresholding techniques [27,31,32] would produce masks that are too coarse to be considered as an adequate initialization for the subsequent CVLSF model, because they frequently exclude large portions of GA regions that cause convergence into local minima during the refinement step. An alternative iterative threshold method based on global image information is proposed here to coarsely segment GA regions from the projection image. This iterative threshold is computed considering a restricted region within the image that gets updated during subsequent iterations, and is set to decrease and converge to a certain value (see later experimental results and analysis section). Assuming that the superscript n indicates the n^{th} iteration, a threshold value Th^n is generated by the OTSU threshold selection method [27], considering the pixels in the original GA projection image I that are within a restricted region NR^n (as indicated in Eq. (1)). This threshold Th^n separates NR^n into foreground and background regions, GR^n and BR^n (indicated in Eq. (2) and (3)), respectively. The updated restricted region in the $n + 1$ th iteration, NR^{n+1} , is obtained by considering the pixels with values under the mean value of the foreground region GR^n (Eq. (4)). We chose the initial restricted region NR^1 to contain all the pixels in I , and a stopping criteria was set for the iteration when the difference between two consecutive thresholds ($Th^{n+1} - Th^n$) is lower than a given constant value (indicated in Eq. (5) set to $c = 0.02$ in our experiments). The foreground region GR^n obtained in the last iteration is then used as the coarsely segmented GA region.

$$Th^n = OTSU(I(i))_{i \in \{NR^n(i)=1\}} \quad (1)$$

$$GR^n(i) = \begin{cases} 1 & I(i) \geq Th^n \\ 0 & I(i) < Th^n \end{cases} \quad (2)$$

$$BR^n(i) = \begin{cases} 1 & I(i) < Th^n \\ 0 & I(i) \geq Th^n \end{cases} \quad (3)$$

$$NR^{n+1}(i) = \begin{cases} 1 & I(i) < \text{mean}(I(k)) \\ & k \in \{GR^n(k)=1\} \\ 0 & I(i) \geq \text{mean}(I(k)) \\ & k \in \{GR^n(k)=1\} \end{cases} \quad (4)$$

$$\text{criterion for next iteration} \begin{cases} \text{continue} & Th^{n+1} - Th^n \geq c \\ \text{stop} & Th^{n+1} - Th^n < c \end{cases} \quad (5)$$

An example of a first iteration is illustrated in Fig. 3. Figure 3(b) shows the histogram of the whole image, background region and GA region corresponding to Fig. 3(a), where manual GA outlines are indicated. The histogram highlights the inhomogeneity of the GA projection image, showing intensity values overlapping between background and GA regions. Figure 3(c) shows that the initial threshold Th^1 (with value shown with a green line in Fig. 3(b)), selected by applying the OTSU method to the whole image, NR^1 , leads to missing a large part of GA regions with similar intensity to the background. Figure 3(d) displays the histogram of NR^1 , with the mean value of the OTSU foreground region GR^1 as new threshold to identify the restricted region for the next iteration, NR^2 . Subsequent iterations are then repeated until the preset stopping criteria are met, producing the coarse GA segmentation displayed in Fig. 3(e). Although the results from this iterative segmentation are coarse and include multiple regions where GA is not present, they proved to serve as a good initial step in the initialization for our methods, where the GA presence is later refined.

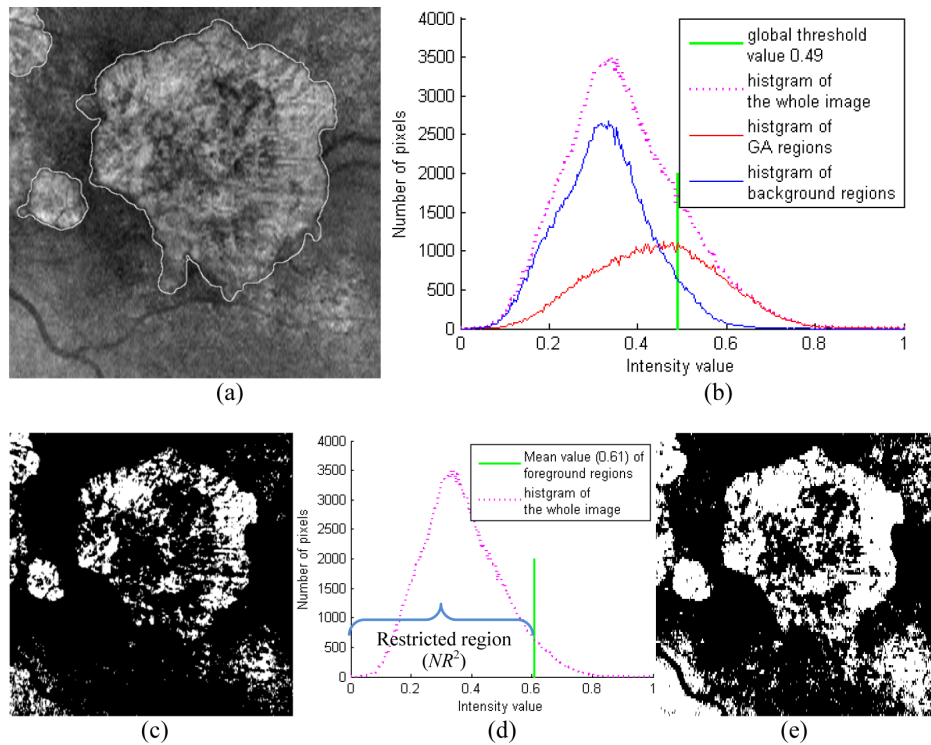


Fig. 3. (a) GA projection image with GA contour generated by manual segmentation. (b) Histogram of GA region and background and the threshold using OTSU method over the whole projection image. (c) Segmentation result obtained by OSTU method in the first iteration. (d) Histogram of the whole image where the mean value of the foreground region resulting from the first iteration and the values corresponding to the restricted region for the second iteration are indicated. (e) Final result for the coarse GA segmentation.

2.3 GA candidate region extraction

The coarsely segmented GA regions obtained in the previous iterative step (section 2.2), are still insufficient to be considered as an initialization outline for the CVLSF method. Isolated low-intensity false negative regions within correctly detected GA regions and extensive false positive regions (as can be observed in Fig. 3(e)) tend to cause “leakage” (segmentation expansion to neighboring structures) in Chan Vese methods, yielding sub-optimal results. A GA candidate region extraction refinement is considered here with the goal of further including isolated background regions and excluding false positive locations in the CVLSF model initialization outline.

Horizontal GA candidate regions are first determined by considering the maximum values of a filtered version of the projection image at each horizontal location, a process illustrated in the example shown in Fig. 4. The generated projection image (example shown in Fig. 4(a)) is first filtered using a moving average filter of 5×5 pixels in order to smooth out possible noise. Figure 4(b) shows the horizontal intensity profile recorded at the vertical coordinate shown in Fig. 4(a) (blue line) and Fig. 4(c) shows the filtered version of this profile. We then compute the maximum value at each horizontal location of the filtered image from the set of vertical scan locations, generating a maximum horizontal intensity profile. Figure 4(d) displays the collection of horizontal filtered intensities recorded at each vertical location in the image, with the maximum horizontal intensity profile shown in Fig. 4(e). The locations in the horizontal axis where the maximum intensity profile takes values higher than a constant threshold (selected experimentally as the mean value of intensity profile set minus its standard deviation) are then considered as horizontal GA candidate locations. The maximum horizontal intensity profile is shown in Fig. 4(f) together with the selected threshold indicated by the red line.

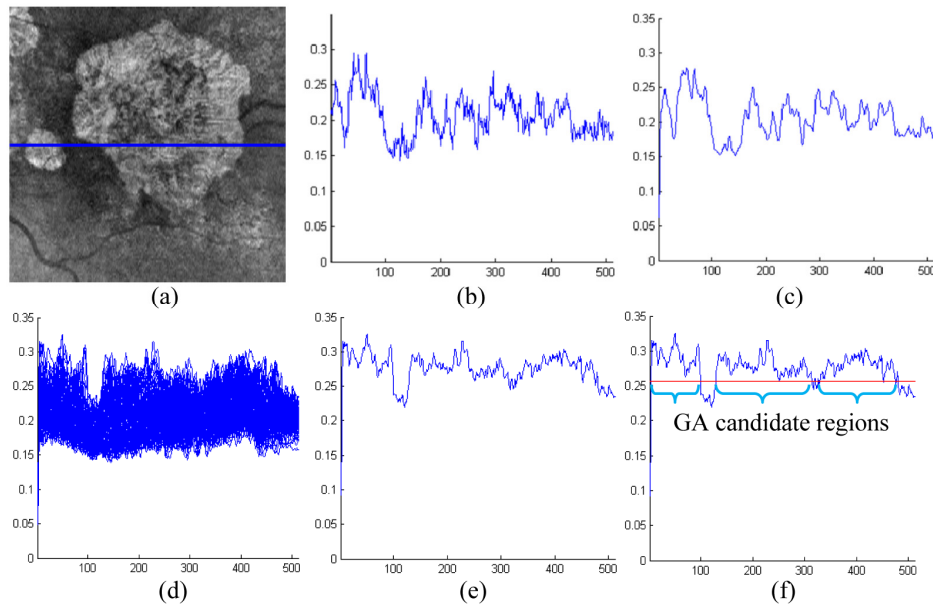


Fig. 4. (a) GA projection image. (b) Projection image intensity profile at the vertical location indicated by the blue line in (a). (c) Intensity profile at the same vertical location after filtering by 5×5 moving average filter. (d) Collection of intensity profiles at all the vertical locations in the image. The maximum horizontal profile is shown in (e) and (f). The red line shown in (e) indicates the selected threshold and the light blue braces indicate horizontal GA candidate regions.

The result obtained from this horizontal candidate region selection is employed to refine the previous coarse GA segmentation (example shown in Fig. 3(e)). These refinements

consist of three consecutive steps: (1) Morphological opening operation applied to remove small isolated regions and smooth borders, with a disk-shaped structuring kernel of radius 0.2 mm; (2) Removal of connected regions at the image border and away from the image center. Isolated regions with all pixels at a normalized distance greater than distance threshold (Dth) from the center of the image and intensity threshold less than an intensity threshold (Ith), where Dth and Ith are set to be 2mm and one third of the maximum intensity of the projection image, respectively, are removed from the segmentation. This is done to eliminate the possible regions related to inhomogeneity in the image and optic nerve head regions, which may be partially appearing showing high intensity values; (3) Region filling using the horizontal candidate region selection. The horizontal regions in each vertical axis that belong to those selected as horizontal candidates and are surrounded by regions selected after the two previous steps are filled to form the final GA candidate region mask. An example illustration of these three refinement steps is shown in Fig. 5.

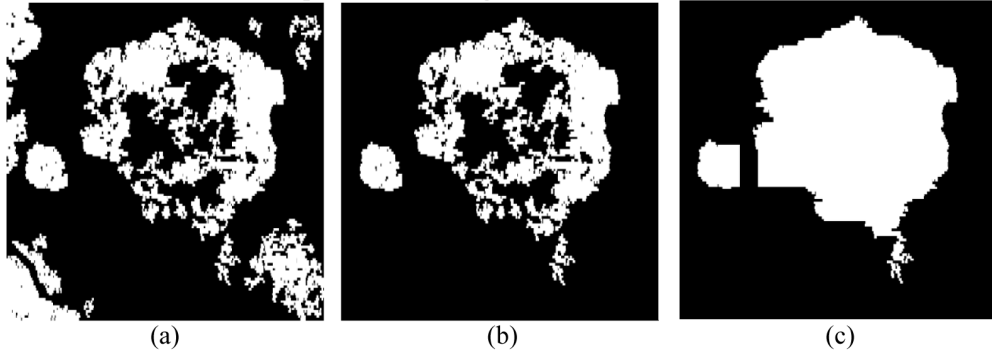


Fig. 5. (a) Result obtained by morphological opening of the coarse segmentation shown in Fig. 3(e), (b) Results after elimination of connected regions at the image border. (c) Refinement results after filling considering the horizontal GA candidate regions.

2.4 Segmentation of GA regions based in an improved C-V model via local similarity factor

The results obtained after the coarse segmentation refinement are taken as an initialization for an improved region-based C-V model [24] with a local similarity factor (CVLSF), which is introduced here to suppress noise influence, while guaranteeing detail preservation in the segmentation results. The objective function for partitioning an image $I(x, y) \in \Omega$ into two regions (GA region and background) is defined as:

$$E(c_1, c_2, C) = \lambda_1 \iint_{in(C)} \left(|I(x, y) - c_1|^2 + LSF_1(x, y) \right) dx dy + \lambda_2 \iint_{out(C)} \left(|I(x, y) - c_2|^2 + LSF_2(x, y) \right) dx dy + \mu Length(C) \quad (6)$$

where $\mu \geq 0$ is fixed constant parameter, and $\lambda_1 > 0, \lambda_2 > 0$ control the contributions of the internal energy and external energy terms, respectively, where object regions taken as internal term are the inside of the contour C ($in(C)$) and background regions considered as external term are the outside of C ($out(C)$). Using the level set definition [28] to represent C , that is, C is the zero level set of a level set function $\Phi(x, y)$, we can rewrite this objective function as:

$$\begin{aligned}
E(c_1, c_2, \Phi(x, y)) &= \lambda_1 \iint_{\Omega} \left(|I(x, y) - c_1|^2 H(\Phi(x, y)) + LSF_1(x, y) H(\Phi(x, y)) \right) dx dy + \\
&\lambda_2 \iint_{\Omega} \left(|I(x, y) - c_2|^2 (1 - H(\Phi(x, y))) + LSF_2(x, y) (1 - H(\Phi(x, y))) \right) dx dy \\
&+ \mu \iint_{\Omega} \delta(\Phi(x, y)) |\nabla \Phi(x, y)| dx dy
\end{aligned} \quad (7)$$

where $H(\Phi(x, y))$ and $\delta(\Phi(x, y))$ are Heaviside function and Dirac function, respectively, which are generally defined as:

$$H_{\varepsilon}(z) = \frac{1}{2} \left(1 + \frac{2}{\pi} \arctan \frac{z}{\varepsilon} \right) \text{ and } \delta_{\varepsilon}(z) = \frac{1}{\pi} \frac{\varepsilon}{\varepsilon^2 + z^2}, z \in R \quad (8)$$

If we keep $\Phi(x, y)$ fixed and minimize the energy function (7) with respect to the constants c_1 and c_2

$$\begin{cases} c_1(\Phi(x, y)) = \frac{\iint_{\Omega} I(x, y) H(\Phi(x, y)) dx dy}{\iint_{\Omega} H(\Phi(x, y)) dx dy} \\ c_2(\Phi(x, y)) = \frac{\iint_{\Omega} I(x, y) (1 - H(\Phi(x, y))) dx dy}{\iint_{\Omega} (1 - H(\Phi(x, y))) dx dy} \end{cases} \quad (9)$$

The local similarity factor is then defined as:

$$\begin{cases} LSF_1(x, y) = \sum_{(i, j) \in N_{(x, y)}} \frac{|I(i, j) - c_1|^2}{d_{(x, y), (i, j)}} \\ LSF_2(x, y) = \sum_{(i, j) \in N_{(x, y)}} \frac{|I(i, j) - c_2|^2}{d_{(x, y), (i, j)}} \end{cases} \quad (10)$$

where $N_{(x, y)}$ represents a neighborhood defined around the central pixel (x, y) (in our experiments defined as a 5×5 pixel window) and $d_{(x, y), (i, j)}$ is the Euclidean distance between pixels located at (x, y) and (i, j) .

Minimizing the energy function (7) with respect to $\Phi(x, y)$, we obtain the corresponding variational level set formulation as follows:

$$\frac{\partial \Phi(x, y)}{\partial t} = \delta(\Phi(x, y)) \left(\lambda_2 |I(x, y) - c_2|^2 - \lambda_1 |I(x, y) - c_1|^2 + \lambda_2 LSF_2(x, y) - \lambda_1 LSF_1(x, y) + \mu \nabla \left(\frac{\nabla \Phi(x, y)}{\|\nabla \Phi(x, y)\|_2} \right) \right) \quad (11)$$

The data term in the CVLSF model (Eq. (11)) is similar to the traditional C-V model [24], differing by the introduction of the local similarity factor LSF. This factor reduces the influence of noisy pixels in the segmentation results. The spatial distance in the LSF was included to balance the intensity difference between neighboring pixels and global mean intensity on two sides of active contour, aiming for convergence to stable value as the number

of iterations increase. The LSF takes into consideration spatial distance and its value penalizes the data term so that the energy function (7) can suppress the influence of noisy pixels. The LSF therefore poses several advantages over the traditional model: (1) Improvement of model against noise of different characteristics. The combination of the spatial distance and gray level difference between a central pixel and surrounding pixels within its local neighborhood balances their similarity and highlights the differentiation between foreground and background regions, reducing noise influence. The LSF value changes in each iteration step, converging to a particular value adapted to the noise properties observed within the local window; (2) Another advantage of the LSF is its ability to preserve more image information. The local spatial information varies adaptively according to the Euclidean distance between neighbor pixels and central pixel, while intensity differences change automatically between the pixels within the local window and the global mean pixel value. We have included a detailed analysis of the method's robustness to noise as compared to the traditional C-V model as well as a comparison of results in GA segmentation in the results section.

In order to accelerate the segmentation speed we also employed coarse-to-fine processing scheme. The GA projection images and initialization outline were first down-sampled to half their resolution in both vertical and horizontal directions and were segmented by the CVLSF method. The results are then up-sampled to the original image resolution and taken as initialization for the fine CVLSF segmentation

2.5 Experimental data and evaluation studies

Our algorithm was implemented in Matlab (The MathWorks, Inc.) and run on a 2.16 GHz Pentium Dual PC with 3 GB memory. Two different data sets were used to evaluate the performance of our algorithm, where all the cases presented with advanced non-nonvascular AMD with extensive GA. Both data sets are the same as described in a previous publication [17]. The first data set consisted of 55 longitudinal SD-OCT cube scans from twelve eyes in eight patients. Each cube consisted of $512 \times 128 \times 1024$ voxels corresponding to a $6 \times 6 \times 2$ mm³ volume in the horizontal, vertical and axial directions, respectively, centered at the macular region of the retina. The second data set consisted of 56 SD-OCT cube scans from 56 eyes in 56 patients with GA. Each cube consisted of $200 \times 200 \times 1024$ voxels corresponding to a $6 \times 6 \times 2$ mm³ volume in the horizontal, vertical and axial directions, respectively, centered at the macular region of the retina. All scans in both data sets were acquired with a Cirrus OCT device (Carl Zeiss Meditec, Inc., Dublin, CA).

For the first data set, manual outlines were drawn by two independent readers in the projection images in two repeated separate sessions, and ground truth segmentation outlines were obtained from them by considering those regions outlined by two or more readers or sessions [17]. For the second data set, manual outlines were available as indicated in corresponding FAF images, which were manually registered to their corresponding location in the projection images and considered as ground truth segmentation. Both data sets were segmented using the method proposed here and a previously proposed method [17] (the later indicated as "QC" method). For the second data set we also had available segmentations as produced by an experimental commercial method (available in Cirrus OCT devices, Carl Zeiss Meditec, Inc., Dublin, CA), which we will call "commercial software". The segmentation results obtained by our algorithm, QC's method segmentation and the commercial software results (for the second data set) were compared quantitatively to the ground truth manual segmentations and independent outlines drawn by the two different readers and sessions (for the first data set). Intra-reader agreement and inter-reader agreement were measured by computing the differences between the manual segmentations drawn at the two separated repeated session and by the two different readers in the first data set.

We employed four metrics to perform these quantitative comparisons: Absolute area difference (*AAD*), overlap ratio (*OR*), correlation coefficient (*cc*) and *p*-value in a *U*-test. The *AAD* measures the absolute difference between the GA areas as segmented by two different methods:

$$\overline{AAD}(X;Y) = \frac{1}{N} \sum_{n=1}^N |Area(X_n) - Area(Y_n)| \quad (12)$$

$$std(AAD)(X;Y) = \sqrt{\frac{1}{N} \sum_{n=1}^N (|Area(X_n) - Area(Y_n)| - \overline{AAD}(X;Y))^2} \quad (13)$$

where X_n and Y_n indicate the regions inside the segmented GA contour of scan n , produced by the methods (or graders) X and Y , respectively. N indicates the total number of images in the set. The OR is defined as the percentage of area in which both segmentation methods agree with respect to the presence of GA over the total area in which at least one of the methods detects GA (Jaccard index). The mean OR and standard deviation values are computed across scans in the data sets in the same way as for AAD

$$\overline{OR}(X;Y) = \frac{1}{N} \sum_{n=1}^N \frac{Area(X_n \cap Y_n)}{Area(X_n \cup Y_n)} \quad (14)$$

$$std(OR)(X;Y) = \sqrt{\frac{1}{N} \sum_{n=1}^N \left(\frac{Area(X_n \cap Y_n)}{Area(X_n \cup Y_n)} - \overline{OR}(X;Y) \right)^2} \quad (15)$$

where the operator \cup and \cap indicate union and intersection, respectively. The correlation coefficients were computed using Pearson's linear correlation between the measured areas of GA computed by the segmentation of different methods or readers, measuring the linear dependence using each scan as an observation. We used the p -value in the U -test to measure the possible statistical differences in the area measured between two segmentation methods or readers.

3. Experimental results and analysis

3.1 Parameter analysis

As explained earlier (in section 2.1), the GA projection image is generated considering a depth of sub-RPE parameter h . This parameter indicates the axial lower limit in pixels underneath the RPE, used to create the GA projection images and was selected through experimentation in a separate set of images. We established the value of h so as to maximize GA separability in the projection image, with separability defined as

$$S_{GA} = \frac{\max_{i=1}^{254} \sum_{j=1}^K |B_i(j) - G(j)|}{K} \quad (16)$$

where B_i indicates the binarized result with threshold i , where the GA projection image was previously normalized linearly to take values in the $[0 \ 255]$ range, and G represents a manually obtained GA segmentation gold standard. K denotes the total number of pixels in the GA projection image. Measured this way, separability reflects GA segmentation feasibility assuming a constant threshold and it is therefore related to clarity in the differentiation of GA from background in the image. We measured the average separability in a set of 43 SD-OCT cubes from 10 eyes, separate from the cases included in the following evaluation of our methods, at different values of depth parameter, ranging from 100 to 300 pixels (approximately 0.2~0.6mm) in 10 pixel steps. Figure 6(a) shows the measured average GA separability at different depth of sub-RPE region. We can observe that GA separability increased as the depth parameter increased from 100 to 220 pixels, maintained a stable value from 220 to 240 pixels, and then decreased as the depth value increased from 240 to 300

pixels. We then selected a value of $h = 230$ pixels (approximately 0.45mm) for the generation of the RSVP projection images that are later segmented using the methods introduced here.

Figure 6(b) displays a scatterplot of detected GA area when using our proposed method vs. the actual GA area, evaluating the performance of the proposed method on a variety of images with different degrees of GA extent. The high correlation coefficient ($cc = 0.973$) observed between the proposed method and the manual gold standard indicates that the measurement of GA regions by the proposed method is adequate, given the proposed method parameters.

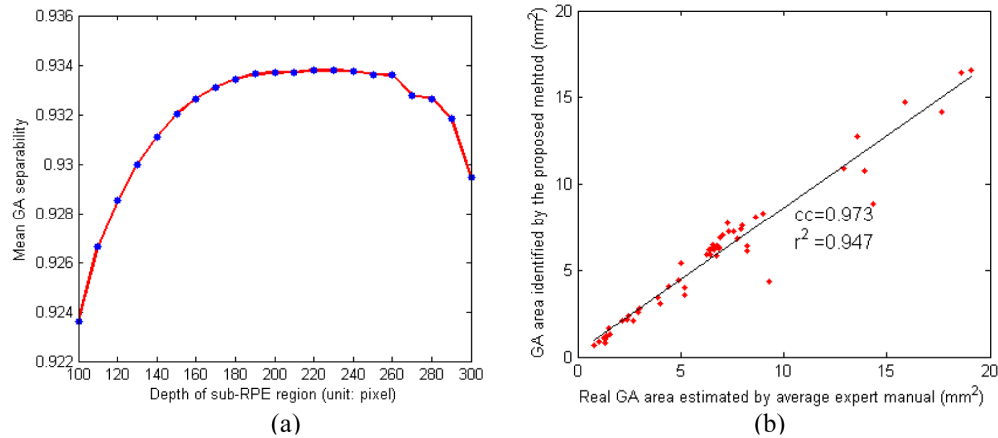


Fig. 6. (a) Average GA reparability across a set of 43 SD-OCT cubes for different values of depth of sub-RPE parameter (h). (b) a scatterplot of estimated area vs actual area of the GA regions.

The iterative threshold method relies on comparing the difference between thresholds at subsequent iterations to a given constant value as a stopping criterion. This value was considered so that the majority of GA regions are included in the coarse results, while ensuring there are only small number of false positive pixels. Figure 7 describes an example that shows two coarse results obtained by different threshold values, as selected by the stopping criterion explained above and as generated when the iterative process converged. The threshold value variation throughout subsequent iterations is also shown. We can observe how the stopping criteria implemented here (Fig. 7(a)) produced a better coarse segmentation than considering the convergence value (Fig. 7(c)).

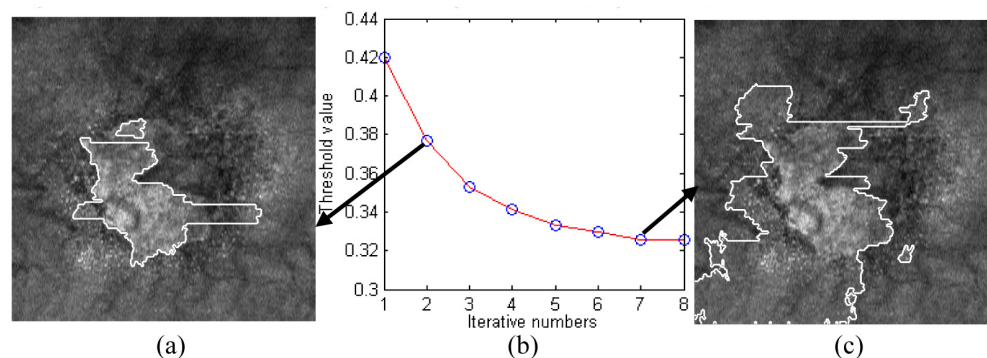


Fig. 7. (a) Refined coarse result with the stopping criteria presented here. (b) Iterative threshold variation throughout subsequent iterations. (c) Refined coarse result at convergence. The white outlines represent the refined coarse results.

3.2 Evaluation of robustness at different noise characteristics

In the CVLSF model, the LSF was introduced to penalize the global data term of C-V model so that the CVLSF model can reduce the influence of the noise, so the CVLSF model can be considered as an improved C-V model rather than a local region-based model. As a consequence, we evaluated the robustness of the CVLSF method in the presence of noise having different characteristics as compared to the traditional C-V model [24] in a synthetic image (of size 61×64 pixels) and in the first data set. The robustness of the CVLSF model mainly depends on contribution of the LSF defined in a local window. The local window size, which specifies the set of neighborhood pixels, controls the extent of noise preservation. The experimental shows that the segmentation accuracy keeps relatively stable for the local window size of 5×5 . In the experiment, a 5×5 local window centered on each pixel is set for the local similarity factor.

Results on the synthetic image are shown in Fig. 8, where parameters in the C-V model and CVLSF model are set to be $\lambda_1 = \lambda_2 = 1$, $\mu = 2500$, $\varepsilon = 0.1$, $\Delta t = 0.1$ and $\lambda_1 = \lambda_2 = 1$, $\mu = 0.05$, $\varepsilon = 0.1$, $\Delta t = 0.5$, respectively. This synthetic image was corrupted by artificial noise of varying known characteristics and strength using the function *imnoise* in *Matlab* to simulate the noise in GA projection image. The uncorrupted image consisted of a background of constant value and three targets with constant intensity different than the background. Two types of random noise realizations were used to corrupt the original image: Gaussian noise and speckle noise. Each noise type was tested at five different noise variance levels. We can observe how the segmentation becomes a harder task as the noise level increases. The C-V model produced segmentation results that were more affected by the noise level than the proposed CVLSF for all noise distribution characteristics, with accuracy decreasing as noise level increased. On the other hand, our CVLSF method presented a lower dependence with noise level, maintaining a relatively similar accuracy at increasing levels of noise.

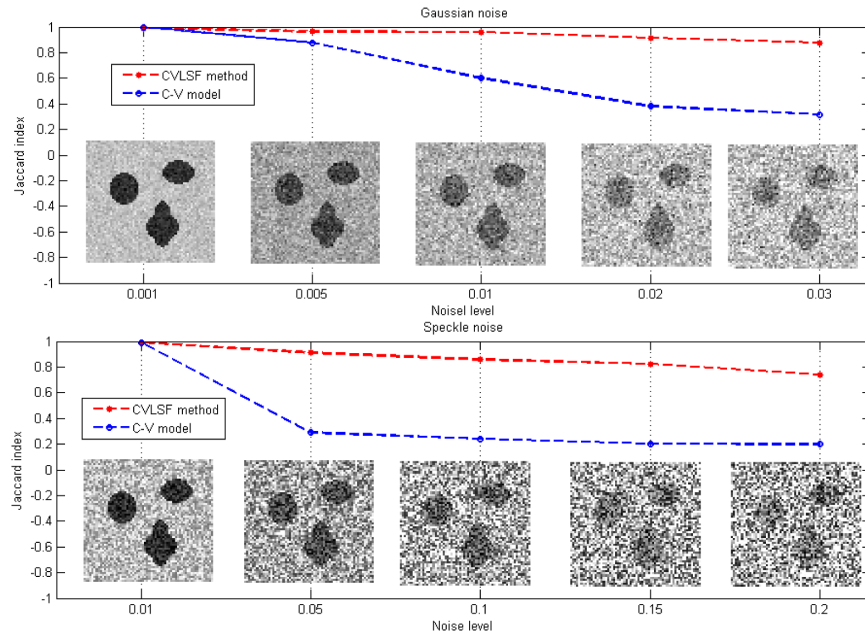


Fig. 8. Quantitative comparison of segmentation accuracy of our proposed CVLSF method (red dashed line) and the traditional C-V method (blue dashed line) at different noise levels. The figure is divided in three rows for two different noise distribution characteristics, with the particular images used for testing displayed in each row. The images in the first row were corrupted by Gaussian noise under 5 noise levels, i.e., $\{0.001, 0.005, 0.01, 0.02, 0.03\}$ and the images in the second row were corrupted by Speckle noise level under 5 noise levels, i.e., $\{0.01, 0.05, 0.1, 0.15, 0.2\}$.

In order to demonstrate the robustness of the CVLSF model in GA projection images, a comparison of the segmentation accuracy results of the traditional C-V method, the traditional C-V method with bilateral filtering as a denoising pre-processing step, and the proposed CVLSF method in the first data set is shown in Fig. 9, where we set $\lambda_1 = \lambda_2 = 1$, $\mu = 2500$, $\varepsilon = 0.1$, $\Delta t = 0.1$ in the C-V model and set $\lambda_1 = \lambda_2 = 1$, $\mu = 0.03$, $\varepsilon = 0.05$, $\Delta t = 0.01$ in the CVLSF model. The figure displays the mean accuracy in terms of Jaccard index (blue bars) in the first data set and the outlines on an example GA projection image resulting from each segmentation method. We can observe that the C-V model produced segmentation results (mean Jaccard index of 0.42 and 0.48 for the non-filtered and filtered versions, respectively) that were more affected by the noise level and intensity inhomogeneity than the proposed CVLSF for original images (mean Jaccard index 0.82). The visual results in an example also demonstrated that the CVLSF model provides better results than C-V model with non-filtered and filtered image. These results in Figs. 8 and 9 suggest the better performance of the proposed method and higher robustness to noise characteristics.

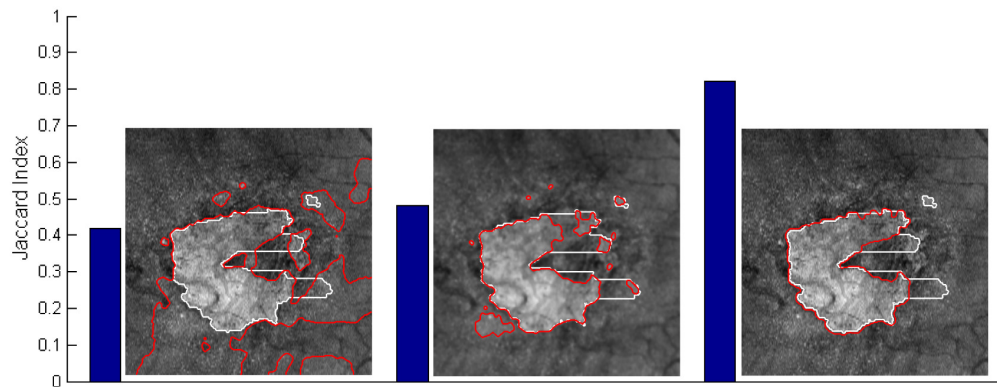


Fig. 9. Quantitative comparison of segmentation accuracy of our proposed CVLSF method and the traditional C-V method in the first data set. The first and second columns show the results of the traditional C-V method and the traditional C-V method with bilateral filtering as a denoising pre-processing step, respectively. The third column shows the results for the proposed CVLSF method. The blue bars indicate the mean Jaccard index resulting from each segmentation method in the first data set, and the images on the right of each bar show the segmented outlines on an example GA projection image. Red and white outlines are the final segmentation results and initialization outlines, respectively.

3.3 Evaluation of GA segmentation

Figure 10 displays several examples with GA regions of different size in the testing data set, where red outlines indicate the segmentation results for the CVLSF model. These examples show cases with different intensity inhomogeneity and complexity, in which accurate GA segmentation is a difficult challenge. We can observe that the outlines produced by the method presented here were relatively precise, given the difficulty of the task.

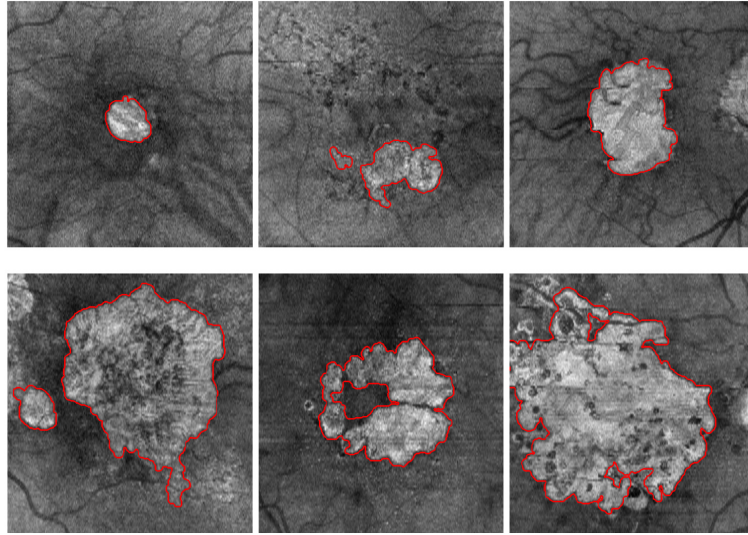


Fig. 10. Examples displaying the automatically segmented GA regions in SD-OCT projection images.

Figure 11 displays example collected manual outlines in examples from the first test data set (as indicated in section 2.5) with outlines made by the two readers at the two repeated sessions indicated with different colors. The intra-observer and inter-observer differences can be visualized. The quantitative results in inter-observer and intra-observer agreement evaluation for this first data set are summarized in Table 1, where A_i ($i = 1, 2$) represents the segmentations of the first grader in the i -th session, and B_i ($i = 1, 2$) represents the segmentations of the second grader in the i -th session. Inter-observer differences were computed by considering the union of both sessions for each grader: $A_{1\&2}$ and $B_{1\&2}$ represent the first and second grader, respectively. The intra-observer and inter-observer comparison showed very high correlations coefficients (cc) and U -test p -values, indicating very high linear correlation and no statistical differences both between different readers and for the same reader at different sessions. The overlap ratios (all $> 90\%$) and the absolute GA area differences (all $< 5\%$) indicate very high inter-observer and intra-observer agreement, highlighting that the measurement and quantification of GA regions in the generated projection images seem effective and feasible.

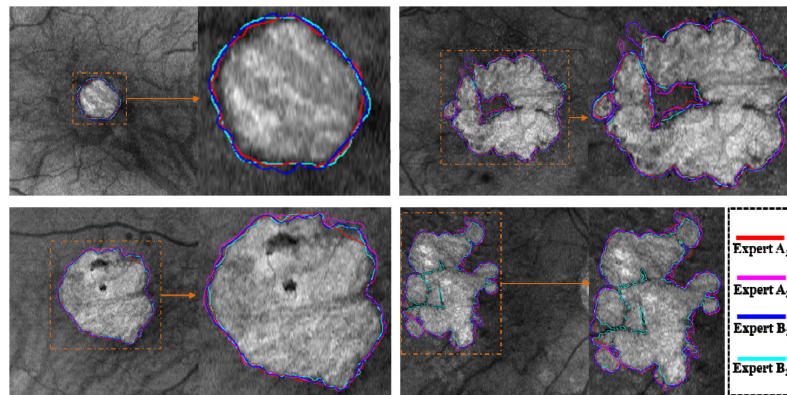


Fig. 11. Manual segmentation examples by two different experts and at two different sessions outlined in RSVP projection images. The region of interest outlined in orange in each RSVP projection image is also shown zoomed in for larger detail. The color label for each observer and session outline is indicated in the legend in the bottom right.

Table 1. Intra-observer and inter-observer correlation coefficients (*cc*), paired *U*-test *p*-values, absolute GA area differences (AAD) and overlap ratio (OR) evaluation

Methods compared	patients/ cubes	<i>cc</i>	<i>p</i> -value (<i>U</i> -test)	AAD [mm ²] (mean, std)	AAD [%] (mean, std)	OR [%] (mean, std)
Expert A ₁ - Expert A ₂	8 / 55	0.998	0.658	0.239 ± 0.210	3.70 ± 2.97	93.29 ± 3.02
Expert B ₁ - Expert B ₂	8 / 55	0.996	0.756	0.243 ± 0.412	3.34 ± 5.37	93.06 ± 5.79
Expert A _{1&2} - Expert B _{1&2}	8 / 55	0.995	0.522	0.314 ± 0.466	4.68 ± 5.70	91.28 ± 6.04

We evaluated the performance of the proposed segmentation algorithm in the first data set by comparing its results to the manual segmentation gold standard and to the previously published QC's method. The results obtained for four example cases are shown in Fig. 12. We can observe that for these cases, the GA outlines obtained by QC's method (yellow outlines) slightly deviate from the gold standard boundary (expert average, in red), whereas the segmentation results obtained by our method (green outlines) seem closer to such gold standard. Table 2 summarizes the results of the quantitative comparison between our algorithm proposed here and manual gold standard (average expert segmentation) and between the previous QC's method and gold standard. The values obtained by our algorithm are displayed in the table in bold face and between parentheses. We also compared the differences of each method to each of the manual readers and sessions independently. Overall, our method presented higher similitudes to the manual gold standard than QC's method, presenting higher correlation coefficients (0.979 vs 0.97), lower absolute area differences (12.95% vs 27.17%), and higher overlap ratio (81.86% vs 72.6%). Lower area differences indicate the area estimated by our method seems closer to the values measured by hand by an average reader than when estimated by the previous method, which would translate into a more accurate GA characterization. The differences observed between our method and the manual gold standard were also very similar to those between our method and each of the independent readers. These differences were higher than the inter-observer and intra-observer differences shown in Table 1, but they were within the same ranges. In fact, the paired *U*-test in measured GA area differences between automated method and each of the manual segmentations was not significant (all with *p*-value > 0.05), while it was significant for differences between QC's method and manual segmentations (all with *p*-value < 0.05). This indicates the results produced by our method seem more similar to manual outlines than QC's method. In conclusion, our algorithm showed better segmentation performance than QC's method when compared to the manual segmentation.

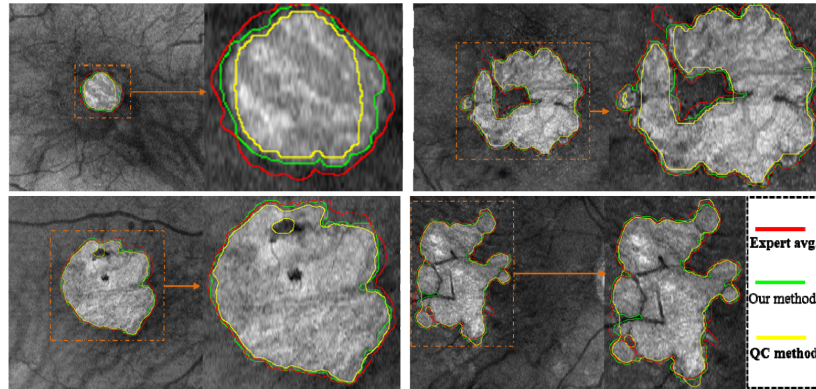


Fig. 12. Segmentation results using the proposed method, QC's method and average expert segmentation (considered as manual gold standard). The cases shown are the same as in Fig. 11 for direct comparison. The region of interest outlined in orange in each RSVP projection image is also shown zoomed in for larger detail. The color label for each segmentation method is indicated in the legend in the bottom right.

Table 2. Quantitative comparison of our algorithm segmentation results (shown in boldface and between parenthesis) and QC's method results to manual gold standard (Avg. Expert) and individual reader segmentation

	QC's (Our Seg.) vs. Avg. Expert	QC's (Our Seg.) vs. Expert A ₁	QC's (Our Seg.) vs. Expert A ₂	QC's (Our Seg.) Vs. Expert B ₁	QC's (Our Seg.) vs. Expert B ₂
patients/ cubes	8 / 55	8 / 55	8 / 55	8 / 55	8 / 55
cc	0.970 (0.979)	0.967 (0.975)	0.964 (0.976)	0.968 (0.976)	0.977 (0.975)
<i>p</i> -value (<i>U</i> -test)	0.026 (0.221)	0.047 (0.389)	0.024 (0.201)	0.017 (0.138)	0.022 (0.191)
AAD [mm ²]	1.438 ± 1.26 (0.811 ± 0.94)	1.308 ± 1.28 (0.758 ± 0.99)	1.404 ± 1.31 (0.853 ± 1.04)	1.597 ± 1.33 (0.984 ± 1.08)	1.465 ± 1.14 (0.897 ± 1.05)
AAD [%]	27.17 ± 22.06 (12.95 ± 11.83)	25.23 ± 22.71 (12.62 ± 12.86)	26.14 ± 21.48 (13.32 ± 12.74)	29.21 ± 22.17 (14.91 ± 12.65)	27.62 ± 20.57 (14.07 ± 11.78)
OR [%]	72.60 ± 15.35 (81.86 ± 12.01)	73.26 ± 15.61 (81.42 ± 12.12)	73.12 ± 15.15 (81.61 ± 12.29)	71.16 ± 15.42 (80.05 ± 13.05)	72.09 ± 14.82 (80.65 ± 12.51)

A set of example results in the second data set evaluated is shown in Fig. 13, where the outlines generated by manual segmentation, commercial software, QC's method, and our method are displayed. We can observe that our method produced results that were similar to the manual outlines, correcting limitations observed in prior methods. Table 3 summarizes the quantitative evaluation in this second data set, comparing each segmentation method (our method presented here, QC's method, and the commercial software) to the manual outlines drawn in FAF images. The correlation coefficients between areas measured using different methods were very high, and all *U*-test *p*-values testing for differences in area measurements showed no statistical significance (*p*-value > 0.05). The overlap ratio was the highest (70%) between our method and the manual segmentation in FAF images, while it was lower than in the previous data set (Table 2), most probably due to the intrinsic differences between SD-OCT and FAF images and possible bias introduced by the registration process. Surprisingly, the differences in AAD between our algorithm and manual segmentations ($1.215 \pm 1.58\text{mm}^2$) are slightly higher than between Qiang's method and manual segmentation ($0.951 \pm 1.28\text{mm}^2$), but both were in the same ranges. The higher overlap ratio with the manual markings observed for our method, but also slightly higher AAD as compared to QC's method, may be due to QC's method producing slight regions or both over- and under-estimation of GA regions, while the method presented here had overall higher similitudes to the manual outlines.

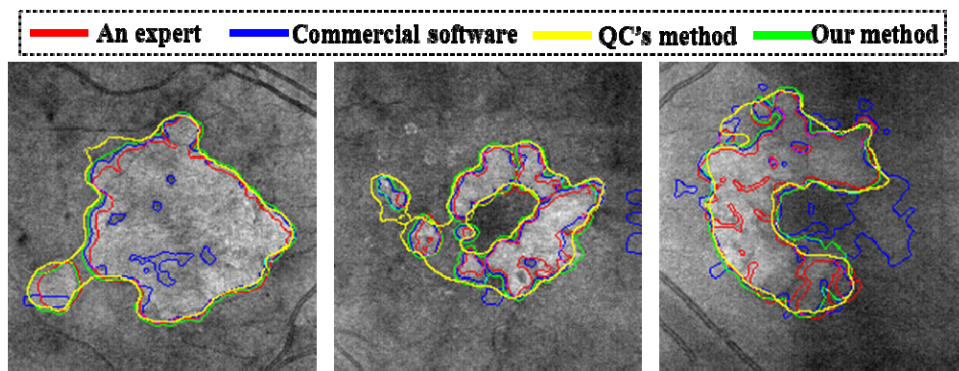


Fig. 13. Comparison of outlines generated by manual segmentation, commercial software, QC's method and our method presented here in three GA patients from the second data set. The color employed for each outline is indicated in the legend on top of the images.

Table 3. Correlation coefficients (cc), paired p-values U-test, absolute differences and overlap ratio in areas of GA between our segmentation method (Our Seg.), QC's method, commercial software segmentation (Com. Sw. Seg.), and expert segmentations manually outlined in FAF images (FAF)

Methods compared	patients / cubes	cc	p-value (U-test)	AAD [mm ²] (mean, std)	AAD [%] (mean, std)	OR [%] (mean, std)
QC's Seg. - FAF	56 / 56	0.955	0.524	0.951 ± 1.28	19.68 ± 22.75	65.88 ± 18.38
Our Seg. - FAF	56 / 56	0.937	0.261	1.215 ± 1.58	22.96 ± 21.74	70.00 ± 15.63
Com. Sw. Seg. - FAF	56 / 56	0.807	0.140	1.796 ± 2.51	34.13 ± 38.62	62.40 ± 21.16

4. Discussion

In this work, we have presented a novel automated GA region segmentation method in SD-OCT images. As summarized in Table 2, our method demonstrated very high accuracy when compared to a manual gold standard generated by two different readers and repeated at two separated sessions (mean OR = 81.86% ± 12.01%; AAD = 0.811 ± 0.94mm²; cc = 0.979; U-test *p*-value = 0.221), and also higher than another known semi-automated technique [17]. Our method also showed good agreement with manual segmentations drawn in FAF images and later registered to the OCT image domain, presenting higher overlap than for particular commercial software and the prior semi-automated technique (Table 3). The example images shown in Fig. 12 and Fig. 13 corroborate these findings, highlighting the similitudes between our proposed segmentation method and manually drawn outlines. We anticipate that the robust results produced by our method may aid the automated characterization of GA area, extent, and location, providing a quantitative, objective and reliable approach to measure and track GA expansion and progression of advanced non-exudative age-related macular degeneration (AMD). These quantifications may aid the discovery of GA-related imaging biomarkers and the development of techniques for the prediction of GA appearance and extension using automated analysis of OCT images, in a similar manner as currently investigated for advanced exudative AMD [30]. Compared to the previously published QC's method [17] and other known semi-automated methods like the one described in [18], our method has several advantages. The main advantage is the fully automated nature of our method, reducing time and possible subjective judgement derived from observer interaction while also producing high-accuracy results, enabling its application in large image data sets.

One of the main difficulties in automated GA segmentation in SD-OCT images is the high noise level and variability, as image quality and noise characteristics vary throughout images acquired using machines from the same vendor and even more so across different vendors. A key aspect of our work to overcome this difficulty is the design of an improved Chan-Vese method considering a local similarity factor (CVLSF). The level-set nature of the method allows the algorithm to handle change of topological structure and irregular shapes easily. The introduced local similarity factor (LSF), balancing similarities observed by the spatial distance and gray level differences within a local window, presents properties that allows the results to be less sensitive to noise of higher intensity and of different characteristics. In section 3.2 we compared the results at known noise characteristics for the CVLSF method and the traditional C-V model, resulting in a higher accuracy observed for the CVLSF model for all conditions tested (Figs. 8 and 9). Another novel aspect of the proposed method is the initialization of the CVLSF segmentation. We designed an iterative threshold segmentation method and a candidate region refinement technique to generate this initialization.

On the other hand, the method described here still has some limitations. Regions of high intensity pixels recorded in the choroid sub-volume considered in the projection image that do not correspond to GA may produce artifacts in the method CVLSF initialization that may remain in the final segmentation results. These regions of over-segmentation could be avoided with manual initialization. For the low intensity regions caused by the deep choroidal

vessels and drusen, we can improve GA projection image by considering the contribution beneath the RPE layer to fill the low intensity regions produced by the presence of choroidal vessels or remove the artifacts by calculating the maximum horizontal intensity profile of the GA projection image at each horizontal location to refine the low intensity regions. We adapted the second strategy to reduce the effect from the deep large vessels in our work. Although the iterative threshold segmentation and refinement using GA candidate regions produced overall adequate initialization results, they may not be optimal for some SD-OCT images presenting complex choroidal structures. In the future, we plan to integrate more GA associated characteristics derived from a three-dimensional approach in the initialization, which may further improve the segmentation results and reduce the sensitivity of our method to initialization. Another limitation of our algorithm is that removal of the region belonging to the optic nerve head, which normally presents high intensity values in the choroid sub-volume, relies on the Euclidean distance from its coordinates to the image center. Although these process seems robust eliminating such regions from the segmentation results, it may also cause removal of potential isolated GA regions that are located far from the macula center, on the edges of the projection image. For example, Fig. 5(a) displays an isolated GA region on the upper left corner of the image that is removed as part of this refinement process, as shown in Fig. 5(b). Another limitation of this process is that a priori assumption that the SD-OCT image is centered at the center of the fovea. We are currently working in methods for detection of optic nerve head within macular SD-OCT images as well as methods to accurately center the image in a pre-processing step in order to reduce their influence of this limitation while also preserving isolated GA regions at the image borders.

5. Conclusions

This paper presents a novel algorithm for automated GA segmentation in SD-OCT images to enable robust, accurate, and objective quantitative measurements of GA extent and location automatically. The proposed method combines a region-based C-V model with a local similarity factor in projection images of a choroid sub-volume. This technique seems more robust to presence of noise, while preserving image detail. Quantitative experimental results demonstrate that the algorithm shows good agreement when compared to manual segmentation by different experts at different sessions and to a consensus manual gold standard, resulting in higher agreement than with a previously known semi-automated method and a commercially-available software package. The proposed algorithm may be clinically useful in providing relatively reliable GA quantitative data that may improve tracking of GA extent, location and expansion in patients diagnosed with advanced non-exudative AMD.

Acknowledgments

This work was supported by China Scholarship Council, 201406840031; the six talent peaks project in Jiangsu Province (2014-SWXY-024); the fundamental research funds for the Central Universities under Grant No. 30920140111004, Qing Lan project under Grant No. D040201; and a Spectrum-SPADA innovation grant of Stanford University, part of the Clinical and Translational Science Award (CTSA) program, funded by the National Center for Advancing Translational Sciences (Grant: UL1 TR001085) at the National Institutes of Health (NIH).